DOCUMENT RESUME

ED 281 167 CS 008 776

AUTHOR Valencia, Sheila; Pearson, P. David
TITLE New Models for Reading Assessment. Reading Education No. 71.
INSTITUTION Bolt, Beranek and Newman, Inc., Cambridge, Mass.; Illinois Univ., Urbana. Center for the Study of Reading.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Dec 86
CONTRACT 400-81-0030
NOTE 14p.
PUB TYPE Reports - Research/Technical (143) -- Viewpoints (120)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Educational Theories; Elementary Education; Evaluation Criteria; *Informal Reading Inventories; *Models; *Reading Achievement; *Reading Diagnosis; Reading Processes; Reading Research; Reading Skills; Reading Strategies; *Reading Tests; Student Evaluation; *Testing Problems; Test Validity; Theory Practice Relationship
IDENTIFIERS Test Content

ABSTRACT

During the last 20 years, the relationship between reading instruction and assessment has been framed by the logic of mastery learning. The ultimate goal of mastery learning is assuring a given achievement outcome across students by varying input characteristics such as the amount and kind of instruction and practice, or the attention paid to prerequisite skills and aptitude characteristics of individual learners. Although the focus on testing has increased with the call for educational reform, currently available tests do not reflect advances in reading theory, practice, or research. The best possible assessment of reading seems to occur when teachers observe and interact with students as they read, evaluating the way in which the students orchestrate resources to construct meaning. An effective framework for complete reading assessment takes into consideration the relationship between objectives, decision-making units, and methods of assessment, and contains three critical features: (1) the attributes should reflect a theoretically sound model of the reading process, (2) the attributes should be highly interdependent and cannot be measured discretely, and (3) whatever is worthy of assessment ought to be assessable in different contexts for different purposes using a variety of strategies. In refocusing statewide assessment of reading, novel concepts that encourage strategic reading and that redefine the assessment-instruction link are being evaluated, such as summary writing, metacognitive judgments, question selection, multiple acceptable responses, and topic familiarity. The goal for reading instructors should be to develop valid, reliable, and usable strategies for assessment of reading comprehension. (Twenty-one references, a list of contrasts between new views of reading and current practices in reading assessment, and a framework of the relationship between objectives, decision-making units, and assessment methods are included.) (NKA)

ED281167

CENTER FOR THE STUDY OF READING

Reading Education No. 71

NEW MODELS FOR READING ASSESSMENT

Sheila Valencia
&
P. David Pearson

University of Illinois at Urbana-Champaign

December 1986

University of Illinois
at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

Bolt Beranek and Newman Inc.
10 Moulton Street
Cambridge, Massachusetts 02238

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

The work upon which this publication is based was performed pursuant to Contract No. 400-81-0030 of the National Institute of Education. It does not, however, necessarily reflect the views of this agency.

CS008776

Abstract

Recent research in reading has reshaped our understanding of the reading process. At the same time, with the call for educational reform, there has been an increased focus on testing. Yet currently available tests do not reflect the advances in reading theory, practice, or research. They are based upon one view of the reading process, while the best reading instruction in our schools is based on an alternative and contradictory point of view. As the result, there is a serious discrepancy between what we know and what we measure. The authors set the stage for possible solutions, showing the relationship between goals, decision making units, and methods of assessment. They propose a new framework for educators to use to develop sound assessment strategies that are based upon a strategic model of the reading process.

New Models for Reading Assessment

Reading assessment has not kept pace with advances in reading theory, research, or practice. As a result, we find ourselves in a dilemma. On the one hand, we argue vehemently for richer and more liberating instructional materials and practices--practices that help students become sophisticated readers, readers who have a sense of ownership and awareness of their reading habits and strategies. On the other hand, we stand idly by and observe yet another round of standardized or end-of-unit basal tests. Even those of us who argue that the current tests do not measure what we mean by reading secretly (sometimes publicly) take pride that our pet instructional technique produces greater gains in learning than another technique on one of those very tests. In this paper, we explain the nature, contributing factors, and consequences of this frightening dilemma in order to set the stage for some possible solutions.

The accountability movement of the 1970's, the wave of recent national reports (Education Commission of the States, 1983) and the focus of the effective schools research (Fisher, Berliner, Filby, Marliave, Cahan, Dishaw, & Moore, 1978) have set the stage for major educational reforms. In most instances, these reports have had reading achievement as one major focus, and in many cases they have relied on students' standardized test scores as measures of effectiveness or educational quality. Such a reliance has led to an increased focus on testing: minimal competency testing, norm-referenced, and criterion-referenced testing.

As evidence of the increasing use of tests, one need only to point out that there are presently at least 40 statewide competency testing programs in place. Add to this the thousands of locally regulated testing programs, the criterion-referenced tests accompanying every basal reading program, and the countless school- and teacher-made tests, and the picture of a nation of schools, teachers, and students engulfed by tests is complete. No matter the perspective one takes on this picture, the conclusion is inescapable: the influence of testing is greater now than at any time in the history of schooling.

The time has come to change the way we assess reading. The advances of the last 15-20 years in our knowledge of basic processes in reading have begun to impact instructional research (Pearson, 1985), and there are hints that they are beginning to find a home in instructional materials and classroom practice (Pearson, 1986). Yet the tests used to monitor the ability and achievement of individual students and to make policy decisions at the school, district, and state level have remained remarkably impervious to advances in reading research (Farr & Carey, 1986; Johnston, in press; Pearson & Dunning, 1985). If we are ever to witness the full impact of this new research upon instruction--if we are to foster a healthy link between assessment and classroom instruction--we must develop new measures of reading.

New Views of the Reading Process

One of the major contributions of recent research has been to articulate a strategic view of the process of reading (e.g., Collins, Brown, & Larkin, 1980; Pearson & Spiro, 1980). This view emphasizes the active role of readers as they use the clues printed on the page to "construct" a model of the meaning of the text. It deemphasizes the notion that progress toward expert reading is guided by the aggregation of component reading skills. Instead, it suggests that at all levels of sophistication, from kindergartener to research scientist, readers use available resources (e.g., text, prior knowledge, environmental clues, and potential helpers) in order to make sense of the text at hand. Progress toward expert reading is guided by the increasing awareness and sensitivity of readers to issues of how, when, and why those resources can best be used. This strategic view also suggests that skilled, but not unskilled, readers can use knowledge flexibly--they can apply what they have learned from reading to new situations (e.g., Campione & Brown, 1985; Spiro & Meyers, 1984); in fact, the ability to use knowledge flexibly predicts how well students will acquire future knowledge.

We have asserted that assessment has not been touched by this strategic view of reading. Figure 1 describes the litany of conflicts we see between what is known about reading and what is done to assess it.

[Insert Figure 1 about here.]

The point is simple but insidious: As long as reading research and instructional innovations are based upon one view of the reading process, while reading assessment instruments in our schools are based upon an alternative and contradictory point of view, we are likely to nurture tension and confusion among those charged with the dual responsibility of improving instruction and monitoring student achievement.

This tension could easily transform itself into a kind of schizophrenia among reading program directors and reading teachers. While anxious to implement instructional practices based upon the latest research, they are plagued by the threat of low test scores. As a result, they are forced to try to integrate two diametrically opposed curricula--one based upon what is measured by the tests for which they are accountable and one based upon what they have learned from recent research.

The Relationship Between Instruction and Assessment

The instruction/assessment link has been vigorously debated by educators. Some sing the praises of instructional programs driven by test results (Haney, 1985; Popham & Rankin, 1981)--or, what some have called outcome-based education--and cite positive results to support their case (e.g., Popham, Cruse, Rankin, Sandifer, & Williams, 1985). Opponents of such testing schemes argue that tests should follow, rather than lead, curriculum (e.g., Berlak, 1985). They claim that overreliance on test scores leads to a narrowing of the curriculum, a tendency to teach to the test, and an emphasis on lower level, more easily tested skills (e.g., Linn, 1985). Still others (e.g., Madaus, 1985) remind us that some of our large-scale tests have become so generic and curriculum insensitive that they are virtually useless for making decisions in a school setting.

The current relationship. During the last 20 years, the relationship between assessment and instruction in reading curriculum has been framed by the logic of mastery learning, introduced in the early 1960s and developed fully by the end of the decade. The goal in mastery learning is to assure a given achievement outcome across students by varying input characteristics such as the amount and kind of instruction, the amount of practice, or the attention paid to presumably prerequisite skills or to aptitude characteristics of individual learners (the age-old notion of matching the method to the child).

The mastery learning systems for teaching reading that evolved in the 1970s tended to focus upon only one tangential feature--a preference for monitoring complex behaviors in component steps or subskills. As a result, the focus in reading mastery systems was to provide extra practice in the form of additional worksheet activities for very specific skills. This practice resulted in a model of reading instruction which emphasized a series of discrete, enabling skills (such as specific letter sound correspondences and specific comprehension skills like sequencing events or locating main ideas). For each enabling skill there was a criterion-referenced test and a set of worksheets upon which to practice the skill. By the late 1970s virtually every major basal reader publisher had its own series-specific management system. In the most recent development, testing companies have begun to offer the option of criterion subskill scores for the items in their norm-referenced tests.

What has happened, of course, is that by conceptualizing reading as the mastery of small, separate enabling skills, there has been a great temptation to operationalize "skilled reading" as an aggregation--not even an integration--of all these skills; and, of course, "instruction" becomes a

succession of opportunities for students to practice these discrete skills on a series of worksheets, workbook pages, and ditto sheets.

The essence of the assessment/instruction relationship under a mastery learning framework is this: Mastery learning encourages us to think of reading instruction as a matter of making certain that students master a "scope and sequence" of enabling skills. The underlying metaphor is that of an assembly line: The "reader" moves along that assembly line, picking up a new part (new skill) at each station along the way. When all the parts are in place, we have a reader ready to tackle real reading. Or do we?

Some hidden dangers. There are some serious consequences of the discrepancy between using one model to define skilled reading and the assessment-instruction link and another to define reading assessment. One danger lies in the false sense of security we are apt to feel if we equate skilled reading with high test scores on our current batch of reading tests. A close inspection of the tasks students are asked to complete in these tests would cast doubt upon any such conclusion.

A second danger stems from the potential insensitivity of current tests to changes in instruction motivated by strategic views of reading. A group of teachers, administrators and policy-makers bold enough to establish a new program might abandon it as ineffective on the basis of a no, or only a small, measurable advantage over a conventional program. They might never consider the alternative interpretation: that the tests they are using are insensitive to effective instruction.

A third danger relates to the issue of curriculum change. Given the strong influence of assessment on curriculum, we are likely to see little, if any, change in reading instruction without an overhaul in our tests. Conscientious teachers want their students to succeed on reading tests; not surprisingly, they look to these tests as guides for instruction. In the best tradition of schooling in Western Civilization, they teach to the test, either directly or indirectly. Tests that portray an inappropriate model of skilled reading will foster inappropriate instruction. Even worse, this situation may lead to the schizophrenic reading curriculum problem we discussed earlier.

A fourth danger stems from the lure of objectivity associated with commercially published tests and the corollory taint of subjectivity associated with informal assessment. For whatever reasons, teachers are taught (and apparently learn) that the data from either standardized or end-of-unit basal tests are somehow more trustworthy than the data that they collect each day as a part of the normal course of teaching. The price we pay for such a lesson is high, for it reduces the likelihood that teachers will use their own data for decision-making within their classrooms.

An alternative relationship. Consider a completely different relationship between instruction and assessment in reading--one where assessment and instruction are synonymous. Based upon a strategic view of the reading process, every act of reading and assessment requires the orchestration of the many resources available in the learning environment--including the text, the reader's prior knowledge, other learners, and the constraints of the situation itself--to construct a satisfactory model of meaning. The goal of every act of reading, and therefore every act of assessment, is identical, regardless of who is performing it. What varies across readers, situations, and levels of sophistication is exactly how readers orchestrate available resources.

Given such a view, the best possible assessment of reading would seem to occur when teachers observe and interact with students as they read authentic texts for genuine purposes. As teachers interact with students, they evaluate the way in which the students orchestrate resources to construct meaning, intervening to provide support or suggestions when the students appear on the verge of faltering in their attempt to build a reasonable model of the meaning of the text. This model, referred to as dynamic assessment (Campione & Brown, 1985), emanates from Vygotsky's notion of the "zone of proximal development," that region of development just far enough--but not too far--beyond the students' current level of competence such that sensitive teachers, using scaffolding tools

such as modelling, hints, leading questions and cooperative task completion, can assist learners in moving to their next level of sophistication. Instruction consists, in such a model, not of remediating deficient skills, but of using assessment strategies and observation to determine which of the potentially useful resources students have trouble using to their best advantage and then providing support and guidance in its application. The "measure" of students' ability is not a score; instead, it is an index of the type and amount of support required to advance learning.

This scenario in which there is no difference between reading, instruction, and assessment is an ideal. While this model is one that may never be fully integrated into large-scale tests of reading, it holds enormous promise for classroom and individual student assessment.

A View to the Near Future

What we must do, then, is to help educators and policy-makers at all levels begin to think about assessment strategies that are consistent with strategic reading and that redefine the instruction-assessment link. What we need are not just new and better tests. What we need is a new framework for thinking about assessment, a framework in which educators begin the process by considering types of decisions they need to make and the level of impact of those decisions.

The framework we propose for developing a complete assessment system is depicted in Figure 2. The attributes of skilled reading listed in the first column represent current working hypotheses about what it means to be a good reader--those outcomes which students and teachers agree are desirable. Columns 2, 3, and 4 indicate levels of impact (or really aggregations of individuals) at which various types of decisions are made. The assumption behind the framework is that different sorts of decisions will have to be made at each level and that each type of decision may require, or lend itself to, different kinds of data and different types of assessment strategies. There are three critical features of the framework:

1. Attributes listed in the first column must reflect a theoretically sound model of the reading process (those listed happen to be our current candidates based upon our reading of the research on basic reading processes).

2. These attributes are highly interdependent and cannot be measured discretely. In fact, a constructive, interactive model of the reading process dictates that the skilled reader selectively chooses from many skills and then orchestrates the combination to produce meaning.

3. Whatever one decides is worthy of assessment ought to be assessable in different contexts for different purposes using a variety of strategies. That is, broad goals at the state level must be consistent with those at the district, school, classroom and individual level. By contrast, assessment of instructional techniques, daily lesson goals and the like belong at the classroom level. The "big goals" cross all levels of decision-making. The smaller goals (the means) do not.

[Insert Figure 2 about here.]

For example, because of the large number of students that come under their respective "responsibilities," units as large as federal and state education agencies, and even most large school districts, need some measures that can be administered to large groups of students to determine general trends. However useful these tools may be for estimating trends, because they sample such a limited range of the repertoire of achievements of any individual student, they are likely to provide invalid and unreliable information for making decisions about individuals.

Some decisions made about individuals may simply require one-on-one interview techniques. In trying to validate our paper and pencil measures of reading (see page 7 for examples), we have conducted hundreds of interviews with individual children who have answered these novel machine-scorable items; one of the lessons of this type of validation effort is that there are some things you can learn only by talking to a student one-on-one. And it could be that student-constructed responses of the type usually generated on daily assignments are most useful in making decisions about groups or whole classes of individuals.

For example, students' sensitivity to the demands of the task, audience and situation and their ability to vary reading strategies to meet these demands (metacognitive strategies) might best be assessed by observing and interacting with students while they are actually applying these strategies in "real" reading situations (Palincsar & Brown, 1984; Palincsar & Brown, 1986). As we illustrate later, we can and should measure these skills in formats amenable to large-scale assessment. But there will always be some limitations to data gathered from group tests of metacognitive activities: (1) what students say may differ from what they do, (2) strategic readers may be too flexible and adaptive to allow us to capture their skill in a small sample of situations and options, and (3) for many readers, these strategies operate at an unconscious, automatic level inaccessible to verbalization or even reflection. In short, here is a case in which large-scale assessment may prove moderately useful for some very limited purposes and decisions; however, the assessment strategies that really count are likely to occur at the classroom or individual level.

It should not be inferred that we are advocating machine-scorable formats for large-scale testing and constructed response formats for individual assessment. There are possibilities for both formats at different levels of impact. While the constraints of large-scale assessment have historically lead to multiple choice, machine-scorable tests, there is some indication that this tradition is changing; in fact, the writing assessment administered by the National Assessment of Educational Progress, along with those administered by most states, require students to write compositions that have to be evaluated by human judges. This change in writing assessment reflects the advances in our understanding of the writing process. Those of us in reading can take at least one lesson from the writing field: The professionals in that field stood firm in demanding large-scale assessments in writing that exhibited face, curricular, and instructional validity until they got them.

Conversely, we can see the usefulness of some machine-scorable formats at the individual assessment level if these items can be constructed to provide opportunities for students that are usually reserved for student-constructed responses.

We suspect that there will be machine-scorable formats useful at the individual assessment level and open-ended formats useful and necessary at the district, state or national levels. Nonetheless, if something like the framework we advocate in Figure 2 ever does come to pass, we predict that, on average, as one moves from large-scale to school to classroom to individual levels, assessment will become more informal, open-ended, and frequent. In fact, the dynamic assessment scenario we presented earlier could become a norm at the reading group level.

In the same way as the mode of assessment is likely to change from large-scale to individual student assessment, some of the outcomes may differ from the state to the classroom level. For example, assessment of instructional techniques, daily lessons and the like belong at the classroom level. These help the teacher and principal understand how instruction may be altered to provide for maximum learning. This assessment of "the means," however, must never take priority over the larger goals of constructing meaning. To do so would place us right back at the point of isolated skills instruction and assessment. The point is, the goals of reading assessment should be consistent across all levels of decision-making, the assessment of the means takes place more frequently, more informally and more appropriately in the classroom.

A Call for Action

Because we are aware that standardized, norm-referenced tests are still the most prevalent type of testing in United States' schools, one clearly important focus for our immediate research should be to develop and evaluate new assessment techniques that are both consistent with our understanding of reading and reading instruction and amenable to large scale testing. Unless we can influence the shape of large-scale assessment, we may not be able to refocus assessment at all. To this end, we have undertaken, with our colleagues around the State of Illinois, an effort to reshape statewide assessment of reading We are working with formats that lend themselves to large-scale assessment and concepts that encourage strategic reading and that redefine the assessment-instruction link. In our pilot work, with approximately 15,000 students in grades 3, 6, 8 and 10, we are evaluating many novel formats:

*Summary writing. Students read 3 or 4 summaries written by other students in response to the selection they just read. They pick the summary they think is best or that the teacher will think is best. In one version, students are presented with a list of features of summaries and check off the reasons for their choice.

*Metacognitive judgments. Students encounter a scenario about a task they might have to perform in response to the selection they have just read (retelling it to different audiences--a peer, a younger child, and a teacher). Then they rate the helpfulness of several different responses for each audience, purpose, or context.

*Question selection. From a set of 20 possible questions, students pick the set of 10 that they think will help a peer best understand the important ideas about a selection.

*Multiple acceptable responses. One of the most productive aspects of a discussion is that a group can consider many alternative candidates as acceptable responses to good questions, especially inferential or evaluative questions. In one format, students select as many responses as they think are plausible. In another, students grade four or five responses on a three-point scale ranging from really complete to on the right track to totally off-base (in much the way a teacher grades short answer or essay responses).

*Topic familiarity. We have developed two machine-scorable formats for assessing prior knowledge. In one, students predict (on a yes-no-maybe scale) whether certain ideas are likely to be included in a selection about a specified topic or theme. In another, they rate the relatedness of vocabulary terms to a central concept of the selection, for example, human blood circulation.

We hasten to add that we are not alone in our effort to influence large-scale assessment by creating tests that reflect these new goals of strategic reading and a redefinition of the assessment-instruction link. For example, the state of Michigan has been involved in a similar effort for the last three years (Wixson, Peters, Weber, & Roeber, 1987). At least a dozen states and several basal publishers have expressed interest in redefining their assessment frameworks.

However, while researchers begin to explore the theoretical and psychometric aspects of new large group formats and techniques, we must all work to develop the disposition and specific assessment techniques needed to answer the varied questions posed by people charged with decisions at different

levels. Our goal, as reading eductors, should be to develop valid, reliable, and usable strategies to be included in all three columns of Figure 2. Only if we are able to fill all the cells with conceptually sound assessment strategies will we approach our goal of equipping educators with a portfolio of assessment strategies that they can use to fit the types of decisions they have to make in the personal, educational, social, and political contexts in which those decisions are implemented.

While we have argued that there is an urgency about these issues now that we have not felt in previous eras, we would not want to leave readers with the impression that the battle is completely new. In fact, we have located a similar concern nearly a century old; in 1892 H. G. Wells bemoaned the influence of the external examiners in determining the curricula of secondary schools in England:

> The examiner pipes and the teacher must dance--and the examiner sticks to the old tune. If the educational reformers really wish the dance altered they must turn their attention from the dancers to the musicians. (p. 382)

While some amongst us will argue that musical instruments ought to be outlawed, others of us will strive to change the tune, perhaps in a way that reflects less concern for melody in favor of greater concern for harmony.

References

Berlak, H. (1985). Testing in a democracy. *Educational Leadership, 43,* 16-17.

Campione, J. C., & Brown, A. (1985). *Dynamic assessment: One approach and some initial data* (Tech. Rep. No. 361). Urbana: University of Illinois, Center for the Study of Reading.

Collins, A., Brown, J. S., & Larkin, K. M. (1980). Inference in text understanding. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension.* Hillsdale, NJ: Erlbaum.

Education Commission of the States. (1983). *Calls for educational reform: A summary of major reports.* Denver, CO: ECS.

Farr, R., & Carey, R. F. (1986). *Reading: What can be measured?* Newark, DE: International Reading Association.

Fisher, C. W., Berliner, D., Filby, N., Marliave, R., Cahen, L., Dishaw, M., & Moore, J. (1978). *Teaching and learning in elementary schools: A summary of the beginning teacher evaluation study.* San Francisco, CA: Far West Regional Laboratory for Educational Research and Development.

Haney, W. (1985). Making testing more educational. *Educational Leadership, 43,* 4-13.

Johnston, P. (in press). *Steps toward a more naturalistic approach to the assessment of the reading process.*

Linn, R. (1985). Standards and expectations: The role of testing (Summary). *Proceedings of a National Forum on Educational Reform.* New York: The College Board.

Madaus, G. F. (1985). Test scores as administrative mechanisms in educational policy. *Phi Delta Kappan, 66,* 611-618.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction, 2,* 117-175.

Palincsar, A. S., & Brown, A. L. (1986). Interactive teaching to promote independent learning from text. *The Reading Teacher, 39,* 771-777.

Pearson, P. D. & Spiro, R. J. (1980). Toward a theory of reading comprehension. *Topics in Language Disorders, 1,* 71-88.

Pearson, P. D. (1985). Changing the face of reading comprehension instruction. *The Reading Teacher, 38,* 724-738.

Pearson, P. D. & Dunning, D. B. (1985). The impact of assessment on reading instruction. *Illinois Reading Council Journal, 3,* 18-19.

Pearson, P. D. (1986). Twenty years of research in reading comprehension. In T. E. Raphael (Ed.), *The context of school-based literacy.* New York: Random House.

Popham, W. J., Cruse, K. L., Rankin, S. C., Sandifer, P. D., & Williams, P. L. (1985). Measurement driven instruction: It's on the road. *Phi Delta Kappan, 55,* 628-634.

Popham, W. J., & Rankin, S. C. (1981). Minimum competency tests spur instructional improvement. *Phi Delta Kappan, 62,* 637-639.

Spiro, R. J., & Meyers, A. (1984). Individual differences and underlying cognitive processes. In *Handbook of Reading Research* (pp. 471-504). New York: Longman.

Wells, H. G. (1892). On the true lever of education. *Educational Review, 4,* 380-385.

Wixson, K. K., Peters, C. W., Weber, E. M., & Roeber, E. D. (1987). New directions in statewide reading assessment. *The Reading Teacher, 40,* 749-755.

Hallmarks of a Good Reader	State or District	School or Classroom	Classroom or Individual
Good Readers.....			
use prior knowledge to help them construct meaning from text.			
draw inferences at the word, sentence, paragraph and text levels.			
provide many plausible responses to questions about a text.			
vary reading strategies to fit the text and the reading situation.			
synthesize information within and across texts.			
ask good questions about text.			
exhibit positive attitudes toward reading.			
integrate many skills to produce an understanding of text.			
are fluent.			
use knowledge flexibly.			

Figure 2. The relationship between objectives, decision-making units, and methods of assessment.

New views of the reading process tell us that . . .	Yet when we assess reading comprehension, we . . .
Prior knowledge is an important determinant of reading comprehension.	Mask any relationship between prior knowledge and reading comprehension by using many short passages about a wide variety of topics.
A complete story or text has structural and topical integrity.	Use short texts that seldom, if ever, approximate the structural and topical integrity found in an authentic text.
Inference is an essential part of the process of comprehending units as small as words and as large as complete texts.	Rely predominantly on literal comprehension test items.
The diversity in prior knowledge across individuals as well as the multiplicity of causal relations in human experiences invite many possible inferences to "fit" a text or a question.	Use multiple choice items with only one correct answer, even when many of the responses listed might, under certain conditions, be rendered plausible.
The ability to vary reading strategies to fit the text and the situation is one hallmark of an expert reader.	Seldom assess how and when students vary the strategies they use when the purpose, context and/or text change.
The ability to synthesize information from various parts of the text and across different texts is important to a complete understanding of the entire text.	Rarely go beyond finding the main idea of a paragraph or passage.
The ability to ask good questions of text, as well as to answer them, is another hallmark of an expert reader.	Seldom ask students to create or select questions about a selection they may have just read.
All aspects of a reader's experience, including reading and writing experiences that arise from school and home settings, influence reading comprehension.	Rarely view information about literacy experiences with the same degree of importance as information about reading performance.
Reading involves the orchestration of many skills that complement one another in a variety of ways.	Use tests that fragment reading into isolated skills and report performance on each.
Skilled readers are fluent; they have reached a level of word identification processing sufficiently automatic to allow most cognitive resources to be used for comprehension.	Rarely consider fluency as an index of skilled reading.
Learning from text involves the restructuring, application and flexible use of knowledge in new situations.	Often ask the reader to respond to the declarative knowledge presented in the text rather than to apply that information to near and far transfer tasks.

Figure 1. A set of contrasts between new views of reading and current practices in assessing reading.